EXHIBIT A

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 278, No. 8, Issue of February 21, pp. 6521–6531, 2003
Printed in U.S.A.

Sequence Identification and Characterization of Human Carnosinase and a Closely Related Non-specific Dipeptidase*

Received for publication, September 23, 2002, and in revised form, December 4, 2002
Published, JBC Papers in Press, December 6, 2002, DOI 10.1074/jbc.M209764200

Michael Teufel‡§, Vladimir Saudek¶‖, Jean-Pierre Ledig, Annie Bernhardt‡, Sylviane Boularand‡‡, Alexandra Carreau‡‡, Nigel J. Cairns§§, Christopher Carter‡‡, David J. Cowley‡, Danielle Duverger‡‡, Axel J. Ganzhorn‡, Chantal Guenet‡, Blanche Heintzelmann‡, Veronique Laucher‡, Claude Sauvage‡‡, and Tatiana Smirnova‡‡**

*From the ‡Department of Exploratory Research, Sanofi Synthelabo Recherche, 16 Rue d'Ankara, F-67080 Strasbourg, France, ‡‡Department of Neuroscience, Sanofi Synthelabo Recherche, 31 Av. Paul Vaillant Couturier, F-92200 Bagneux, France, ¶Aventis Pharma, Paris Research Center, 13 Quai Jules Guesde, F-94400 Vitry-sur-Seine, France, **Laboratoires Fournier, Département Exploratoire, 50 rue de Dijon, F-21121 Daix, France, and §§Department of Neuropathology, Institute of Psychiatry, Kings College, De Crespigny Park, London SE5 8AF, United Kingdom*

Carnosine (β-alanyl-L-histidine) and homocarnosine (γ-aminobutyric acid-L-histidine) are two naturally occurring dipeptides with potential neuroprotective and neurotransmitter functions in the brain. Peptidase activities degrading both carnosine and homocarnosine have been described previously, but the genes linked to these activities were unknown. Here we present the identification of two novel cDNAs named CN1 and CN2 coding for two proteins of 56.8 and 52.7 kDa and their classification as members of the M20 metalloprotease family. Whereas human CN1 mRNA and protein are brain-specific, CN2 codes for a ubiquitous protein. In contrast, expression of the mouse and rat CN1 orthologues was detectable only in kidney. The recombinant CN1 and CN2 proteins were expressed in Chinese hamster ovary cells and purified to homogeneity. CN1 was identified as a homodimeric dipeptidase with a narrow substrate specificity for Xaa-His dipeptides including those with Xaa = βAla (carnosine, K_m 1.2 mM), *N*-methyl βAla, Ala, Gly, and γ-aminobutyric acid (homocarnosine, K_m 200 μM), an isoelectric point of pH 4.5, and maximal activity at pH 8.5. CN2 protein is a dipeptidase not limited to Xaa-His dipeptides, requires Mn^{2+} for full activity, and is sensitive to inhibition by bestatin (IC_{50} 7 nM). This enzyme does not degrade homocarnosine and hydrolyzes carnosine only at alkaline pH with an optimum at pH 9.5. Based on their substrate specificity and biophysical and biochemical properties CN1 was identified as human carnosinase (EC 3.4.13.20), whereas CN2 corresponds to the cytosolic nonspecific dipeptidase (EC 3.4.13.18).

The dipeptide carnosine (β-alanyl-L-histidine) was first isolated in 1900 from meat extracts (1) and subsequently found to be widely distributed in excitable central and peripheral vertebrate tissues (for review see Ref. 2). This compound is the archetype of a variety of aminoacyl histidine dipeptides such as homocarnosine (γ-amino-butyryl-histidine, GABA[1]-His) or anserine (β-alanyl-L-1-methylhistidine). Whereas the role of many of the carnosine-related dipeptides is poorly understood, the function of carnosine has been studied intensively. In peripheral tissues this compound can be found at high levels (*e.g.* muscle tissues, 1–20 mM) (3) suggesting a crucial role as cytosolic buffer to neutralize lactic acid (4) (for review see Ref. 5). In the central nervous system carnosine meets many criteria for a neurotransmitter modulating synaptic processes but also appears to be involved in neuroprotection (6). The dipeptide has antioxidant and free radical scavenger properties, through complexation of transition metals such as zinc or copper (7, 8); it retards senescence of cultured fibroblasts (9) and is an antiglycation agent (for review see Ref. 10). Homocarnosine was suggested to be a precursor for the neurotransmitter GABA. Being controlled by one or several carnosinases, it acts as a GABA reservoir and may mediate the antiseizure effects of GABAergic therapies (11, 12).

Carnosine is synthesized by carnosine synthase (EC 6.3.2.11; see Ref. 13) from β-alanine and histidine in many tissues and degraded by intra- or extracellular dipeptidases, also named carnosinases, all belonging to the large family of metalloproteases. A cytosolic form previously named tissue carnosinase (EC 3.4.13.18) was first isolated from porcine kidney by Hanson and Smith (14) in 1949 and subsequently found widely distributed in tissues of rodents and higher mammals (15–18). The human isoform purified from kidney was described by Lenney *et al.* (19). The authors suggested, however, that "human tissue carnosinase" acts as a cytosolic nonspecific dipeptidase rather than a selective carnosinase based on its broad substrate specificity and the strong inhibition by bestatin (20). A secreted form of human carnosinase was first described by Perry *et al.* (21) in patients with carnosinemia and was first purified from human placenta (22). The enzyme was also isolated from human plasma and originally named human serum carnosinase (EC 3.4.13.20; see Ref. 23). It was distinguished from its cytosolic counterpart because of its particular distribution in human plasma and brain, its unique capability to degrade homocarnosine, and absence in non-primate mammals except for the Syrian golden hamster (24).

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AX139747 for CN1 and AX523938 for CN2.

§ Contributed equally. To whom correspondence should be addressed. Tel.: 33-388454186; Fax: 33-388459075; E-mail: michael.teufel@sanofi-synthelabo.com.

‖ Contributed equally.

[1] The abbreviations used are: GABA, γ-aminobutyric acid; AGE, advanced glycation end products; CSF, cerebrospinal fluid; EST, expressed sequence tag; MDA, malondialdehyde; OPA, *o*-pthaldialdehyde; CHO, Chinese hamster ovary; HGS, Human Genome Sciences; HMM, hidden Markov model; ORF, open reading frame; HPLC, high pressure liquid chromatography; MALDI-TOF, matrix-assisted laser desorption ionization time-of-flight.

In this study the discovery of two novel genes CN1 and CN2 coding for metallopeptidases of the M20 family is described. We demonstrate that CN1 corresponds to the secreted human carnosinase and CN2 to the cytosolic nonspecific dipeptidase previously named tissue carnosinase.

EXPERIMENTAL PROCEDURES

Materials

The pcDNA3 vector was purchased from Clontech. pCRII was from New England Biolabs. Oligonucleotides were synthesized on an Applied Biosystems DNA/RNA synthesizer model 392. DNA sequence analysis was performed using an Applied Biosystems model 373A DNA sequencer with the Big Dye dyedeoxy terminator cycle reagents (Applied Biosystems, Foster City, CA). Amplification steps using PCR were carried out with a Thermal Cycler 480 (PerkinElmer Life Sciences) using the GeneAmp kit reagents (Applied Biosystems). Heparanized human plasma was obtained from the Centre Regional de Transfusion Sanguine de Strasbourg. LLC-PK1 and CHO-K1 were from the American Type Culture Collection (Manassas, VA). Human serum was obtained from the Centre de Transfusion Strasbourg, France. Protein concentrations were determined using a protein assay kit (Bio-Rad).

SDS-PAGE was carried out using 8% Tris/glycine polyacrylamide gels (Invitrogen). Proteins were visualized by silver staining of the gels using a commercial kit (Silver Express; Invitrogen).

Carnosine and homocarnosine were from Sigma, Ala-Ala, βAla-Ala, Ala-His, Gly-His, Ser-His, Tyr-His, His-His, Glu-His, Ile-His, Met-His, Val-His, *p*Glu-His, βAsp-His, Leu-His, GABA-His, Gly-His-Gly, Gly-Gly-His, Ser-Pro, Ala-Pro, Gly-Gly, and Gly-Leu were from Bachem. *N*-Methylcarnosine was synthesized by Neosystem (Strasbourg, France). Pepstatin, Bestatin, AEBSF, E64, Phosphoramidone, and leupeptin were from Roche Diagnostics. *o*-Pthaldialdehyde 1,10-*o*-phenantrolin, *p*-hydromercurybenzoate, and other chemicals were from Sigma.

Cloning and Sequencing of CN1 and CN2 cDNAs

A particularly abundant set of ESTs with overlapping sequences was identified in the Human Genome Sciences (HGS) human brain cDNA libraries. The 2.0-kb cDNA insert of the EST clone 999021 from this set (shown below to bear the full-length sequence of CN1) was sequenced, and the amino acid structure was deduced using standard algorithms. The insert was excised from pBluescript by *Eco*RI/*Xho*I digestion and subcloned into the cytomegalovirus promoter-based plasmid pcDNA3.1(+) to obtain pCN1. The insert sequence was compared with public and HGS data bases. Another set of homologous ESTs was identified in the HGS data base, and the EST clone 2831791 (bearing CN2) from HGS was sequenced, and the 2.6-kb cDNA insert was subcloned into pcDNA3.1(−) after excision by *Eco*RI/*Bam*HI digestion and subsequently named pCN2.

The 2-kb insert of mouse EST clone AI 746479 carrying the mCN1 gene was identified in the NCBI data base and sequenced, and the amino acid sequence was deduced using standard algorithms. A rat homologue of CN1 (EST clone AA925553) was identified in the GenBank™ data base and sequenced.

Bioinformatic Analysis

Sequence analysis was performed using the Genetic Computer Group Sequence (www.gcg.com). For HMM profiling the CN1, CN2, and orthologues derived from the *Drosophila melanogaster* and *Caenorhabditis elegans*, together with the sequence of a yeast protein with unknown function (Swiss-Prot entry YFL4_YEAST), were used as a seed to build a hidden Markov model profile with HMMer (hmmer.wustl.edu), which was progressively refined during an iterative search of Swiss-Prot and TrEMBL. Fold recognition was carried out with the ProCeryon (www.proceryon.com/) software package.

Site-directed Mutagenesis

Three point mutations, M1, M2, M3, and a double mutant, M3M4, corresponding, respectively, to amino acid changes H133A, D166A, E201A, and E201A+D229A of CN1 ORF, were generated. Mutants were generated by PCR using full-length CN1 cDNA as template and the following oligonucleotides: for M1, 5′-AAA GGC ACC GTG TGC TTC TAC GGC GCT TTG GAC-3′ and 5′-CTC GAG GCG GCC GCT CAT TAG TGA TGG TGA TGG TGA TGG AGC TGG GCC AT-3′; for M2, 5′-ACG GAG GTA GAC GGG AAA CTT TAT GGA CGA GGA GCG ACC GCC AAC AAA-3′ and 5′-CTC GAG GCG GCC GCT CAT TAG TGA TGG TGA TGG TGA TGG AGC TGG GCC AT-3′; for M3, 5′-CTG GAG CAA GAT CTT CCT GTG AAT ATC AAA TTC ATC ATT GAG GGG ATG GAA GCT GCT GGC-3′ and 5′-CTC GAG GCG GCC GCT CAT TAG TGA TGG TGA TGG TGA TGG AGC TGG GCC AT-3′; and for M4, 5′-GCT GTT CCC CCG GGT TCC ATA AGT GAT TGC TGG CTT CCT TTG GCT GAT CCA CAG GTT GGC TGA AAT-3′ and 5′-GAA TTC GTC GAC ATG GAT CCC AAA-3′. In addition, a His_6 domain was incorporated during amplification to follow expression of mutant proteins. Amplified fragments were inserted into pCRII and fully sequenced. Mutated fragments were reintroduced into full-length carnosinase by fragment reassembly. Briefly, M1 mutated fragment was digested by *Dra*III and *Xho*I to obtain 3′ end harboring the mutation (1149 bp) and ligated with the *Eco*RI-*Dra*III CN1 fragment (378 bp) corresponding to the 5′ end. For M2, M2 mutated fragment was digested with *Acc*I to obtain a central fragment containing the mutation (391 bp) that was ligated, together with the *Eco*RI-*Acc*I 5′ end CN1 fragment (463 bp) and with the *Acc*I-*Xho*I 3′ end CN1 fragment (673 bp). For M3, M3 mutated fragment was digested with *Bgl*II and *Xho*I to obtain the 3′ end mutated fragment (969 bp) and ligated to the CN1 *Eco*RI-*Bgl*II fragment. Finally the double mutant M3M4 was obtained after restriction of M3M4 mutated fragment by *Eco*RI and *Sma*I to obtain 5′ end mutated region (731 bp), which was ligated to the *Sma*I-*Xho*I CN1 3′ end. All mutants were inserted into pcDNA3 for expression, and full-length reconstructed mutants were verified by DNA sequencing.

Multiple Tissue Expression Array and Northern Blot

Multiple tissue expression array and Northern blots were from a commercial source (Clontech). Hybridization and washing were performed according to the manufacturer's specifications using ^{32}P-labeled cDNA probes. The probe CN1 was synthesized by PCR with the primers 5′-TTTCAGATAACCTGTGGATCA and 5′-GAACATTTT GGCAATTG-GAATG covering nucleotides 680 to 1374 of the coding sequence. The CN2 probe was synthesized by PCR with the primers 5′-TTCCCTCTT-TCCTTTCCCTC and 5′-GCATACACCACCATGTCTG, located in the 3′-untranslated region.

A commercial multiple mouse tissue Northern blot (Clontech) was hybridized to a 1022-bp ^{32}P-labeled probe obtained from amplification of the mouse CN1 EST clone (accession number AI746749) and the primers 5′-GTGGTGGAGAAACAGGTAAC and 5′-CCAAAGGTTCCT-GAGTGGAA. A commercial multiple rat tissue Northern blot (Clontech) has been hybridized to a 276-bp ^{32}P-labeled probe obtained from amplification of rat cDNA and the primers 5′-TTACACCACCAAGC-CCAATC and 5′-CCTCCCACTCCG TCAGTAAA, designed from the rat EST (accession number AA 925553) containing predicted rat CN1. Hybridization signals were detected with a PhosphorImager (Molecular Dynamics).

Design and Production of Antibodies

Two synthetic peptides containing an N-terminal cystein and CN1 residues 256–272 (peptide $_{256}C17E_{272}$) and 312–329 (peptide $_{312}Y18K_{329}$) were used to raise polyclonal rabbit antibodies (C17E and Y18K) at Eurogentech followed by affinity purification with immobilized antigenic peptides. Purified antibodies were stored at −20 °C in 50% glycerol at a final concentration of 0.5–1 mg/ml. The same protocol was applied to produce antibodies against a synthetic peptide $_{223}S16E_{238}$ (S16E) of human CN2.

Tissue Distribution of CN1 and CN2 Proteins

Immunochemical detection of CN1 and CN2 was carried out by Western blotting of commercial multiple tissue blot (Oncogene, Boston, MA) or of 20 μg of human brain extracts (Clontech), CSF, and plasma using an immunopurified anti-CN1 polyclonal antibody (Y18K 1:1000) and a standard protocol. Tissue distribution of CN2 was studied using a multiple tissue blot (Oncogene) with purified anti-S16E antibody (1:1000). Specific immunocomplexes were visualized using a commercial kit (ECL; Amersham Biosciences).

Immunohistochemistry

Human brain sections of 7 μm were deparaffinized and processed using the microwave antigen retrieval approach in target unmasking fluid buffer (Pharmingen) according to the manufacturer's protocol. The sections were pre-treated with 10% of normal goat serum in phosphate-buffered saline and then incubated with primary antibodies diluted in phosphate-buffered saline/2% normal goat serum for 24–48h at 4 °C. Antibody dilutions were 1:1000 for anti-C17E and 1:500 for anti-Y18K. Specific staining was detected with a commercial kit (ABC; Vector Laboratories) according to the manufacturer's instructions. Slides were examined with a Nikon Diaphot 300 microscope, and images were

captured with an Eastman Kodak Co. digital camera. In control experiments primary antibodies were omitted and replaced by phosphate-buffered saline/2% normal goat serum. The specificity of the anti-C17E antibody was evaluated by pre-incubation with the corresponding antigenic peptide before proceeding to immunostaining of tissue sections. Immunostaining with each antibody was repeated at least twice.

Expression of CN1 and CN2

Transient Transfection of CHO-K1 with Wild-type and Mutant CN1 Genes and CN2 Gene—CHO cell line K1 (ATCC) was cultured according to ATCC instructions. The cells were plated a day before transfection into a 10-cm culture plate (5×10^6 cells per plate). Cells were transfected with 15 μg of pcDNA expression vectors carrying wild-type CN1, M1, M2, M3, or M3M4 or CN2 genes using the LipofectAMINE2000 protocol according to the manufacturer's instructions (Invitrogen). Cells and culture supernatants were harvested after 48 h of incubation in Dulbecco's modified Eagle's medium/H12 medium supplemented with 10% fetal calf serum and analyzed for enzyme activity as described below. The expression of recombinant proteins was verified by Western blot analysis.

Construction of a Stable CHO Cell Line and Large Scale Expression of CN1—A stable CHO-K1 cell line expressing CN1 was constructed using the LipofectAmine 2000 protocol (Invitrogen) according to the manufacturer's instructions. Subclones were obtained by the limiting dilution technique (25). Positive selection was based on the presence of secreted active enzyme and confirmed by Western blot using Y18K antibodies.

Large scale production using the stably expressing CHO cell line was performed with Cell Factory plates (Nunc). 9.5×10^7 cells were plated per unit, amplified for 5 days in Dulbecco's modified Eagle's medium/H12 medium supplemented with 10% fetal calf serum, and incubated for 3 days in serum-free MSSL medium (500 ml per plate), minimum Eagle's medium/HF12 (1:1), 2 mM GlutamaxI, 1% nonessential amino acids, 0.5 mM sodium pyruvate, D-glucose (0.5 g/liter) concentrated lipids (0.1%), ITS supplement (1%), and 0.83 μg/ml $FeSO_4$ (Invitrogen). To the medium, 0.5 mM sodium butyrate was added to further trigger the overexpression of the enzyme by 30% (26). The cell culture medium was concentrated 10-fold by ultrafiltration prior to protein purification.

Determination of Carnosinase Activity

Carnosinase activity was assayed according to a method described by Bando *et al.* (27) modified and adapted to 96-well plates. Briefly, substrate hydrolysis was carried out in 50 mM Tris-HCl buffer, pH 7.5, 1 mM carnosine in a 100-μl final volume using 0.25–0.5 μg of cell extract or 10 ng of purified enzyme. To assay for carnosinase activity in mammalian cells, LLC-PK1 and CHO-K1 cells were cultivated as recommended by ATCC. A soluble cell extract was obtained by centrifugation (15 min at 50000 × *g* at 4 °C) of cell lysates from 10×10^8 sonified cells. Carnosinase activity in cell culture supernatants or heparanized human plasma was determined with a 5-μl sample. The reaction was initiated by addition of substrate and stopped after 60 min of incubation at 30 °C by adding 50 μl of 1% trichloroacetic acid. Liberated histidine was derivatized by adding 50 μl of 5 mg/ml *o*-pthaldialdehyde (OPA) dissolved in 2 M NaOH and 30 min of incubation at 30 °C. Fluorescence was read using a MicroTek plate reader (λ_{Exc}: 360 nm and λ_{Em}: 460 nm). Reaction blank values were obtained by adding the trichloroacetic acid stop solution 1 min prior to substrate addition. Reactions were carried out in triplicate. The activity of CN2 was assayed as described above except that 50 mM Tris-sarcosine, pH 9.5, 0.1 mM $MnCl_2$ was used as buffer.

For the determination of V_{max} and K_m an HPLC-based method, which detects liberated amino acids by precolumn derivatization with OPA, was used (28); 10 μl of enzyme solution (diluted to 1 ng/μl) was added to 50 μl of 100 mM Tris-HCl, pH 7.5, and the volume was adjusted to 90 μl by adding H_2O. The reaction was started by addition of 10 μl of carnosine (0.1–50 mM), homocarnosine, βAla-Ala, Ala-Ala, Pro-Ala, and Ser-Gln, as well as the tripeptides Gly-Gly-His or Gly-His-Gly, and stopped after 60 min of incubation at 30 °C by adding 400 μl of 96% ethanol. After 15 min of centrifugation at 15000 rpm at room temperature, supernatants were transferred to a fresh tube and dried under vacuum using a SpeedVac. The pellet was resuspended in 1 ml of 100 mM sodium borate buffer, pH 9.6. Prior to HPLC separation the samples were derivatized with OPA using an automated method programmed on a HP1100 HPLC system (Agilent, Waldbronn, Germany). Freshly prepared OPA reagent was added 1 min prior to injection, and the derivatized amino acids were separated at 0.3 ml/min on a Kromasil C18 column (2 × 250 mm; Merck) equilibrated in Eluent A (99.5:0.5 of 0.012 M sodium acetate:tetrahydrofuran, pH 6.5) thermostated at 25 °C. A linear gradient was applied from 50–100% amino acid Eluent B (50:35:15 of 0.012 M sodium acetate:methanol:acetonitrile, pH 6.5) for 16 min, and a step elution of 100% amino acid Eluent B was applied for 7 min. Derivatized amino acids were detected by their fluorescence (λ_{Exc}: 340 nm, λ_{Em}: 450 nm). Calculation of kinetic constants was done with MICROCAL ORIGIN™ (Microcal Software Inc., Northampton, MA) using a direct fit of the Michaelis equation to the experimental data.

Purification of CN1

Chromatography on DEAE-Sephacel—Five liters of filtered cell culture supernatant were bound to 500 ml of DEAE-Sephacel (Amersham Biosciences) equilibrated with Buffer A (50 mM Tris-HCl, pH 7.5, containing 0.1 mM $MnCl_2$) and packed into a chromatography column (XK 50/50; Amersham Biosciences). The column was washed at 10 ml/min with 4 column volumes of Buffer A. A linear NaCl gradient from 0 to 500 mM in Buffer A over 3.5 column volumes followed by a step to 1 M NaCl was applied to elute the bound material, and 10-ml fractions were collected.

Size Exclusion Chromatography—Fractions containing carnosinase activity were pooled and concentrated by ultrafiltration on an AMICON YM100 membrane (Millipore, Waltham, MA). The concentrated sample was injected on a Hi-Load Superdex 200 prep grade HR26/60 (Amersham Biosciences) size exclusion column. Proteins were eluted in Buffer A at 3 ml/min, and 3-ml fractions were collected. Pooled fractions were concentrated by ultrafiltration as described above.

Chromatography on MonoQ—The pooled fractions from the size exclusion step were loaded after 2-fold dilution with Buffer A onto a MonoQ HR5/5 column (Amersham Biosciences) equilibrated with the same buffer. Non-specifically bound proteins were eluted applying a step of Buffer A containing 100 mM NaCl followed by a linear gradient from 100 to 200 mM NaCl over 5 column volumes. Then a linear NaCl gradient between 200 and 400 mM in Buffer A over 13 column volumes was applied, followed by a second gradient of 5 column volumes to obtain 500 mM NaCl. Finally a 1 M NaCl step was applied for 5 column volumes.

Purification of CN2

5×10^8 CHO-K1 cells transfected with pCN2 were resuspended in 15 ml of ice-cold resuspension buffer (10 mM Tris-HCl buffer, pH 7.0, containing 100 μM $MnCl_2$), sonicated 5 times for 10 s, and the lysate was cleared by centrifugation at 20000 × *g* for 15 min at 4 °C (S20 fraction).

Chromatography on POROS50 DEAE™—The S20 supernatant was loaded on a POROS50 DEAE™ column equilibrated with resuspension buffer (10 × 100 mm; Applied Biosystems) connected to a BioCAD® work station (Applied Biosystems). Unbound protein was washed with 5 column volumes of resuspension buffer, and the CN2 protein was eluted with a linear gradient from 0 to 500 mM NaCl over 20 column volumes. Fractions of 3 ml were collected, and enzyme activity was assayed with 10 μl of each fraction.

Size Exclusion Chromatography—Fractions containing carnosinase activity were pooled and concentrated by ultrafiltration on an AMICON YM10 membrane (Millipore). The concentrated sample was injected on a Hi-Load Superdex™ 200 prep grade HR26/60 (Amersham Biosciences) size exclusion column equilibrated with Buffer A containing 150 mM NaCl. Proteins were eluted in that buffer at 3 ml/min, and 3-ml fractions were collected. Pooled fractions were concentrated by ultrafiltration as described above.

Chromatography on MonoQ—The pooled fractions from the size exclusion step were loaded after 2-fold dilution with Buffer C (20 mM Tris-HCl, pH 7.5, 0.1 mM $MnCl_2$, 10% (v/v) glycerol) onto a MonoQ column (HR5/5; Amersham Biosciences) equilibrated with the same buffer. Unspecific bound proteins were eluted with Buffer C containing 100 mM NaCl. Then a linear NaCl gradient between 100 and 400 mM NaCl in Buffer C over 20 column volumes was applied, followed by a step of 0.5 M NaCl in Buffer C for 5 column volumes.

Reversed Phase HPLC

The pooled fraction (100 μl) of CN1 or CN2 obtained from the MonoQ purification step were injected on a C4 column (2.1 × 100 mm, Aquabore; Applied Biosystems) equilibrated with 5% acetonitrile containing 0.1% trifluoroacetic acid using the INTEGRAL™ microanalytical work station (Applied Biosystems). The column was washed at 250 μl/min with 20 column volumes of this solvent and then a gradient 5–80% acetonitrile in 0.1% trifluoroacetic acid over 20 column volumes was applied. UV absorption was monitored at 280 nm.

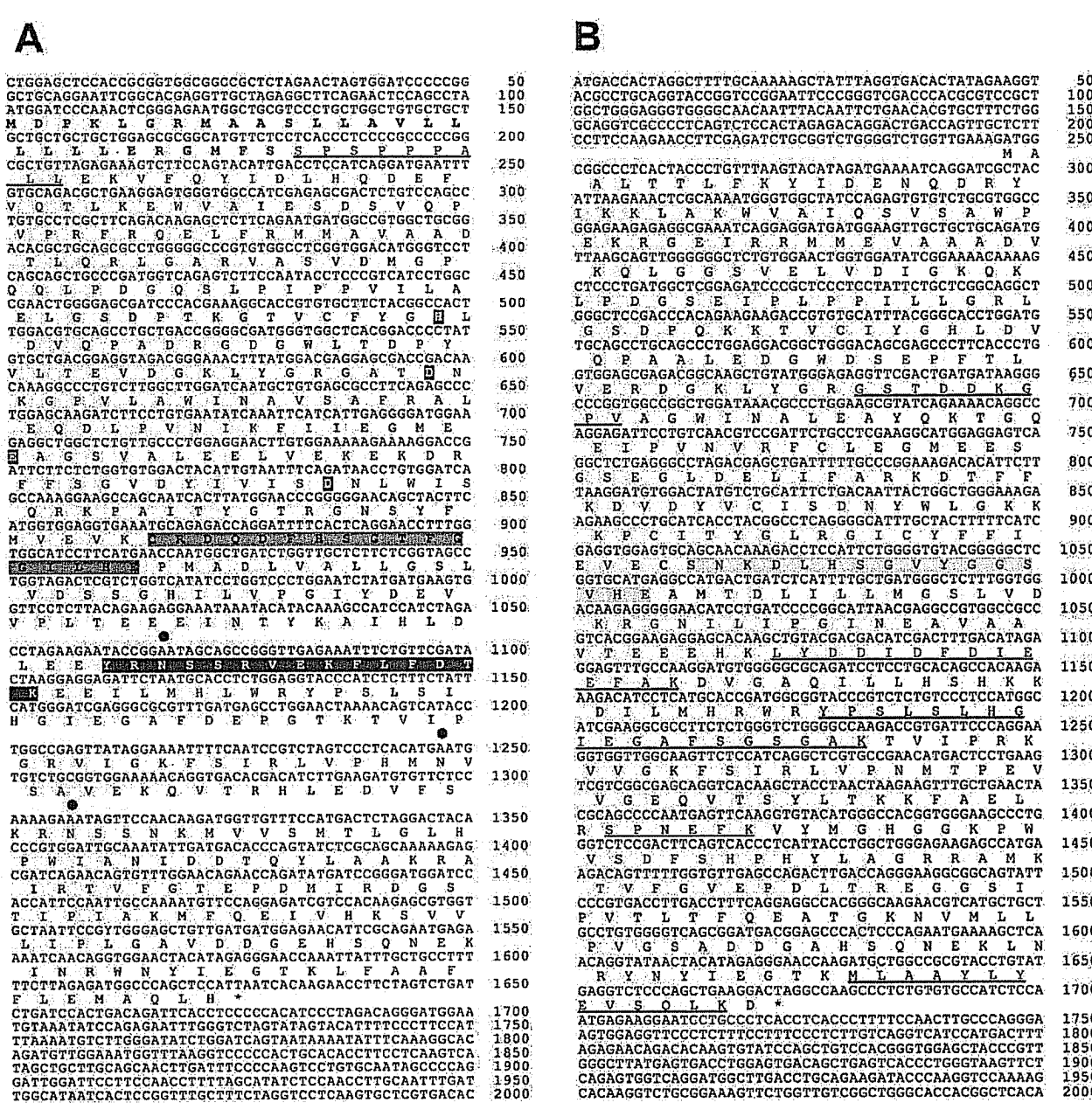

FIG. 1. **Nucleotide sequences and predicted amino acid sequences of cDNAs encoding human CN1 and CN2 proteins.** *A*, CN1, the putative signal sequence is in *bold letters* and the N-terminal amino acids as determined by protein sequencing are *underlined*. Potential *N*-glycosylation sites are marked with *black dots*. *Bold red letters* indicate the H133A, D166A, and E201A/D229A mutations. Sequences used to generate polyclonal antibodies are shown in *colored boxed letters* as follows: *cyan*, C17E; *dark blue*, Y18K. *B*, CN2, the terminal stop codon is marked with an *asterisk*, and the N-terminal amino acid sequences of the purified peptides from LysC digests of the protein are *underlined*. The sequence used to generate polyclonal antibodies (S18E) is shown in *yellow shaded letters*.

MALDI-TOF Mass Spectrometry

MALDI-TOF mass spectrometry was performed on a VOYAGER DEPRO instrument (Applied Biosystems). Matrix (α-cyano-hydroxycinnamic acid or sinapinic acid; Fluka) was freshly prepared in 30% acetonitrile, 0.3% trifluoroacetic acid at a concentration of 25 mg/ml. Crystals containing the protein sample were obtained by the drying droplet method. 1 μl of HPLC fraction was applied to the MALDI sample plate, and 0.5 μl of matrix was added. Spectra were recorded by accumulating 100 shots using the linear mode at 20000-V acceleration voltage, with 1 μs of delayed extraction.

N-terminal Sequencing

Purified CN1 (~300 pmol) were subjected to automated Edman degradation using a 492HT protein sequencer (Applied Biosystems) and standard protocols. The sequence of CN2 was verified using internal peptides obtained from a trypsin digestion of purified enzyme. Peptides were separated by reversed phase HPLC. 100 pmol of purified peptides were subjected to automated Edman degradation as described above.

Analytical Deglycosylation

100 ng of purified CN1 were treated with PNGase F (Oxford Glycosciences) using the following procedure: 200-μl protein samples after MonoQ were dried under vacuum using a SpeedVac. The dried pellet was resuspended in 100 μl of deglycosylation buffer (20 mM $NaPO_4$, 50 mM EDTA, 0.5% SDS, 1 mM β− mercaptoethanol) and then heated to 95 °C for 5 min before 10 μl of Nonidet P-40 were added. The experiment was started by the addition of 30 μl of PNGAseF solution and followed at 37 °C for 24 h. The sample was subsequently analyzed by SDS-PAGE using an 8% polyacrylamide gel (Invitrogen), and proteins were visualized after silver staining of the gel using a commercial kit (Invitrogen).

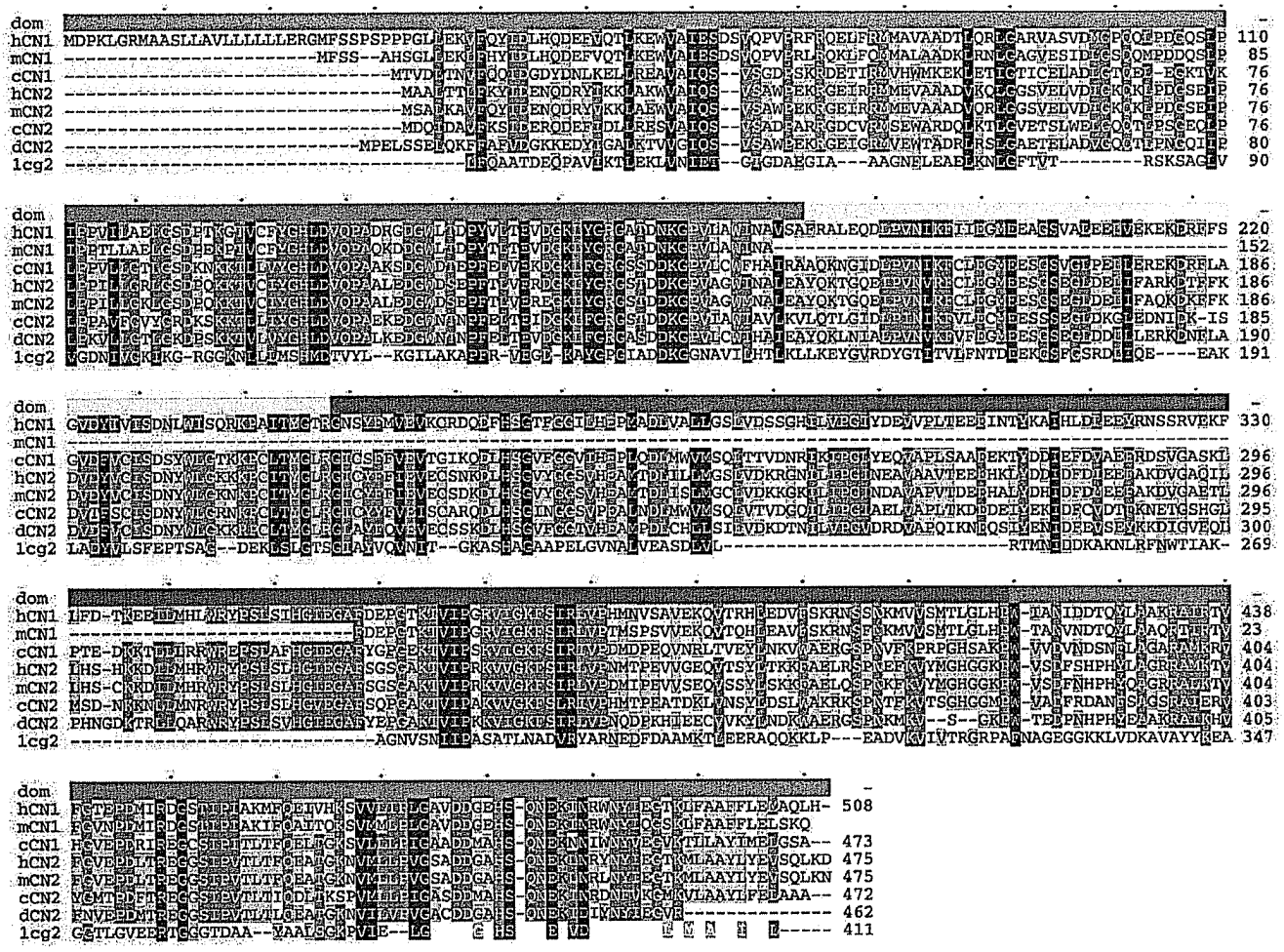

FIG. 2. **Amino acid alignment of human CN1 and CN2 genes and their putative orthologues compared with carboxypeptidase G from *Pseudomonas* sp. of known three-dimensional structure (PDB 1cg2).** Prefixes *h*, *m*, *c*, and *d* stand for human, mouse, *C. elegans*, and *D. melanogaster*. Data base entries are as follows: Swiss-Prot, cCN1, O18000; dCN2, Q9V9H1; GenBank™, dCN2, Y71H2AM (gene predicted in a cosmid using homology and ESTs covering 5′- and 3′- ends AV183130.1, AV195866.1). The active site metal-binding residues are *highlighted* in *red* (histidines) and *yellow* (carboxylic acids). The predicted catalytic domain identified from the 1cg2 structure is indicated in *red*, the 1cg2 dimerization domain is in *blue*, and the CN1 signal peptide is in *green* (determined experimentally). *Shading* in *black*, *dark gray*, and *light gray* indicates 100, 80, and 60% amino acid similarity.

Measurement of MDA-induced Toxicity in SH-SY5Y Neuroblastoma Cells after Transient Transfection with CN1

SH-SY5Y neuroblastoma cells (ATCC) were cultured according to ATCC instructions. The cells were plated a day before transfection into a 10-cm culture plate (5×10^6 cells per plate). Cells were transfected with 15 μg of pCN1 using a Lipofectamine2000 (Invitrogen) protocol according to the manufacturer's instructions. 5 to 6 h after transfection, cells were transferred into 96-well culture plates at a density of 2×10^4 cells/well and cultured for 48 h before treatment with malondialdehyde (MDA). MDA (3 mM) and carnosine treatment were performed as described previously (29). Cell viability was determined 16 h after treatment by assaying mitochondrial dehydrogenases as measured by reduction of 3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide (0.5 mg/ml, 4 h). In control experiments cells were transfected with empty vector pcDNA3.1. Experiments were done in triplicate and repeated five times.

RESULTS

Identification, Cloning, and Sequence Analysis of CN1 and CN2—An assembly of overlapping ESTs (contig) derived from human brain cDNA libraries were identified in the Human Genome Sciences data base, and an open reading frame of 1524 nucleotides could be defined (Fig. 1*A*). One of the clones (EST 999021), containing the full-length sequence coding for an unknown polypeptide of 508 amino acids, was identified and named CN1. The putative ATG translation start codon is in a favorable context for translation initiation (30), and no ATG codons were detected further upstream. Sequence analysis predicted a protein of a molecular mass of 56803 Da with an isoelectric point of pH 4.4, carrying a typical signal peptide sequence and three *N*-glycosylation sites. Further analysis of the HGS data base revealed EST 2831719 from human kidney containing the full-length sequence of a homologue with 49% identity to CN1, named CN2. The ORF of CN2 codes for an unknown cytoplasmic protein of 473 amino acids with a predicted molecular mass of 52872 Da and a pI of pH 5.6 (Fig. 1*B*).

Initially the amino acid sequences deduced from the two full-length cDNA sequences were searched against protein data bases (Swiss-Prot, TrEMBL) and ESTs (GenBank™) using BLAST (31). By this method, homologues of CN1 and CN2 were found among the predicted proteins derived from the genomes of *C. elegans*, *Drosophila melanogaster*, *Saccharomyces cerevisiae*, and homologous ESTs were found for mouse and rat. Interestingly, no orthologue other than human CN1 revealed a typical N-terminal signal peptide. The primary structures for all of the CN orthologues are highly conserved (*e.g.* human CN2 *versus* mouse CN2 was at 91% identity), but the function of these proteins is unknown. Consequently, a progressively refined HMM profile (see Ref. 32) was applied for

The Journal of Biological Chemistry

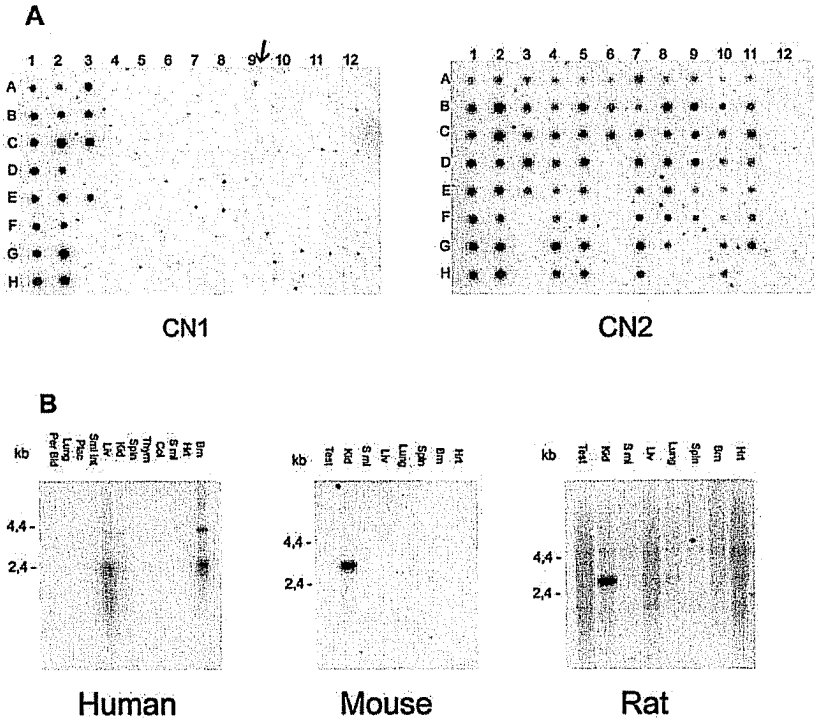

FIG. 3. **Tissue distribution of CN1 and CN2 mRNAs.** *A*, the human multiple tissue expression array (Clontech) was hybridized to a 695-bp ^{32}P-labeled probe corresponding to CN1 ORF 680–1374 (*left panel*) or a 378-bp probe corresponding to the 3′-untranslated region sequence of CN2 (*right panel*). *Arrow*, liver. *B*, Northern blot analysis of CN1 mRNA in human and rodents. *Left panel*, the human Northern blot (Clontech) was hybridized to a 695-bp ^{32}P-labeled probe corresponding to CN1 ORF 680–1374. *Middle panel*, the mouse Northern blot (Clontech) was hybridized to a 1022-bp ^{32}P-labeled probe corresponding to the 3′-untranslated region in EST accession number AI746749. *Right panel*, the rat Northern blot (Clontech) was hybridized to a 276-bp ^{32}P-labeled probe corresponding to the 3′-untranslated region of EST accession number AA925553. Sample annotation on the commercial human multiple tissue expression array is shown in Table I. *Brn*, brain; *Hrt*, heart; *S ml*, smooth muscle; *Col*, column; *Thym*, thymus; *Spln*, spleen; *Kid*, kidney; *Liv*, liver; *Sml int*, small intestine; *Plac*, placenta; *Lung*, lung; *Per Bld*, peripheral blood; *Test*, testis.

data base searches (see "Experimental Procedures"), and a list of sequences with over hundred proteins was retained, all of them being metallohydrolases of the MEROPS MH clan of metallopeptidases (33). This clan contains six families (M18, M20, M25, M28, M40, and M42) with enzymes of known activities such as carboxypeptidases or aminopeptidases. Among the top ranking proteins with known function, a number of enzymes of the MEROPS M20 metallohydrolase family were found, in particular in the M20A subfamily. The essential histidine and carboxyl residues (as shown in Fig. 2) (34) in the metal binding sites of all enzymes of this family, as well as all CN proteins, were found to be completely conserved, providing a strong indication that CN1 and CN2 may be co-catalytic metallopeptidases of the M20 family. To reinforce this hypothesis, a fold recognition search (threading) (see "Experimental Procedures") (35) was performed. With this approach structural homology of CN1 was found with carboxypeptidase G2, a homodimeric metalloprotease (36), which also belongs to the M20A family.[2]

Tissue Distribution of CN1 and CN2—The tissue distribution of CN1 and CN2 mRNA expression was analyzed using a commercial human tissue cDNA array. As shown in Fig. 3*A*, the CN1 gene is expressed in the central nervous system of adults and to a far lesser extent in liver. No expression is visible in the fetal brain. Unlike CN1, CN2 is ubiquitously expressed throughout human tissues (Fig. 3*A*).

CN1 expression in human, mouse, and rat were further studied by Northern RNA analysis (Fig. 3*B*). A hybridization signal of 2.4 kb was observed in human brain and liver, which confirmed the dot blot experiment. A second band of 4.4 kb was visible in human brain mRNA. Although it is likely that this signal derived from contamination of pre-mRNA, our data cannot exclude the possibility that the 4.4-kb band corresponds to a splice variant of CN1. RNA of the mouse and rat orthologues of CN1 was detected exclusively in the kidney and was not found in the brain (Fig. 3*B*).

The presence of CN1 protein in human samples was verified by Western blot (Fig. 4*A*). One or multiple immunoreactive bands were found in extracts of human plasma and CSF (Fig. 4*A*, *lanes 1* and *6*), and two bands were found in cortex and hippocampus (Fig. 4*A*, *lanes 2–4*). Multiple bands suggest the presence of different post-translational modifications of CN1 such as glycosylation. Analysis of a commercial multiple human tissue blot confirmed the absence of CN1 protein in other tissues (not shown). Taken together, our data demonstrate identical patterns of CN1 RNA and protein expression. Expression of the CN2 protein in different human tissues is ubiquitous (Fig. 4*B*) and correlates with the expression pattern of its mRNA (Fig. 3*A*). A strong signal at 54 kDa was found in kidney and liver, and a weaker signal was detectable in brain, spleen, ovary, testis, lung and pancreas, whereas the protein is apparently not expressed in heart.

The presence and cellular localization of CN1 protein in human neurons of temporal cortex and hippocampus was analyzed by immunohistochemistry. CN1 protein was present in the cytosol of pyramidal neurons in the hippocampus and large sized and small sized neurons of the temporal cortex (Fig. 5), whereas in the subcortical white matter small sized cells, presumably glial cells, and neuronal fibers were immunopositive (not shown). All data together suggest constitutive expression of CN1 protein in adult human brain with a fraction of CN1 in the process to be post-translationally modified during the secretion process.

Enzymatic Activities of CN1 and CN2—For functional characterization, CN1 was expressed in transiently transfected CHO cells. The CN1 protein was found in the extracellular medium as demonstrated by Western blotting (see Fig. 7*B*, *lane 1*) confirming the prediction that CN1 is a secreted protein. As described above, the bioinformatic analysis predicted that the CN1 protein belongs to the M20 family of metallohydrolases, comprising homologues of carboxypeptidases, tripeptidases, non-peptidase homologues, and peptidaseV. However, using known substrates for amino- or carboxypeptidases, no amino-

[2] In the course of our work a public Protein Family Database profile (PF01546) became available (www.sanger.ac.uk/cgi-bin/Pfam/getacc?PF01546) that would identify CN1 and CN2 as remote homologues of this profile.

The Journal of Biological Chemistry

The Journal of Biological Chemistry

or carboxypeptidase activity was found in the crude supernatant from cells transfected with CN1 (data not shown).

Subsequently the hypothesis was tested that CN1 codes for the human homologue of the bacterial peptidaseV, which was also identified as a homologue of CN1 by the HMM method. Because this enzyme is also known as βAla-His dipeptidase or carnosinase (E.C. 3.4.13.3), cell culture supernatants, cell extracts, and human plasma were analyzed for carnosine degradation using a fluorescent assay (see "Experimental Procedures"). As shown in Fig. 6, carnosine degradation is only observed in the supernatant of CHO cells transfected with CN1 and in human plasma but not in the cell extracts of transfected cells or with supernatants and cell extracts of non-transfected cells. This is in good agreement with the Western blot data that CN1 protein is a secreted protein.

To test the prediction that CN1 codes for a metalloprotease, three single point mutations removing putative metal binding residues, namely H133A (M1), D166A (M2), E201A (M3), and a double mutant M3M4 (E201A+D229A), were constructed and expressed under similar conditions as for the wild-type gene. Although produced in similar quantities as the wild-type CN1 protein, each of the mutations led to the production of inactive enzyme (data not shown).

As expected, CN2 was found in the cytoplasmic fraction of transiently transfected CHO cells as analyzed by Western blotting (Fig. 7*E*, *lane 1*). Under the assay conditions applied for CN1 (pH 7.5; no addition of metal ions) little cytoplasmic carnosinase activity was detectable; however strong degradation of carnosine was found at pH 9.5 in the presence of 0.1 mM Mn^{2+} (Fig. 6).

Purification and Characterization of Recombinant CN1 and CN2—14 mg of CN1 protein could be purified from 5 liters of cell culture supernatant of stably transfected CHO-K1 cells at a specific activity of 1.3 μmol of histidine formed per min and mg of protein after three chromatographic steps (see "Experimental Procedures"). Purity and homogeneity, as well as absence of degradation products of the enzyme, were assessed by SDS-PAGE (Fig. 7*A*, *lane 4*), Western blot (Fig. 7*B*, *lane 4*), RP-HPLC, and MALDI-TOF mass spectrometry (data not shown). Under denaturing conditions in SDS-PAGE it migrates at ~82 kDa, whereas by MALDI-TOF mass spectrometry the molecular mass of the monomeric subunit was determined to be 63.7 kDa. By both techniques the observed molecular mass for CN1 protein was significantly higher than its predicted molecular mass of 52.8 kDa, indicating post-translational modification of the enzyme. Because three *N*-glycosylation sites are predicted in the protein primary sequence, the recombinant CN1 enzyme was subjected to endoglycosidase treatment. After incubation with PNGaseF, the molecular mass of CN1 protein shifted from 82 kDa (Fig. 7*C*, *lane 1*) to ~61 kDa in SDS-PAGE (Fig. 7*C*, *lane 2*), indicating the removal of carbohydrate moieties. However, three closely migrating bands in SDS-PAGE suggested that the deglycosylation might not have been complete. The molecular mass of native CN1 was 167 kDa as determined by analytical size exclusion chromatography, indicating that the enzyme forms a homodimer in solution.

The N-terminal sequence of CN1 protein was determined by automated Edman degradation. The five N-terminal residues were Ser^{28}-Pro-Ser-Pro-Pro, confirming the removal of the predicted signal peptide (37) at the predicted signal peptidase cleavage site C-terminal to Ser^{27} (38). Using two-dimensional electrophoresis, the isolelectric point was determined at pH 4.5.

CN2 protein was purified by DEAE, size exclusion, and MonoQ chromatography, and 1.3 mg of pure CN2 with a specific activity of 1.0 μmol of histidine formed per min and mg of protein was obtained at pH 9.5 from 1.8×10^9 transiently

TABLE I
Sample annotation on the commercial human multiple tissue expression array

1	2	3	4	5	6	7	8	9	10	11	12
A Whole brain	Cerebel., left	Substantia, nigra	Heart	Esophagus	Colon, transverse	Kidney	Lung	Liver	Leukemia HL-60	Fetal brain	Yeast total RNA
B Cerebral cortex	Cerebellum, right	Accubens nucleus	Aorta	Stomach	Colon, descending	Skeletal muscle	Placenta	Pancreas	HeLa S3	Fetal heart	Yeast tRNA
C Frontal lobe	Corpus callosum	Thalamus	Atrium, left	Duodenum	Rectum	Spleen	Bladder	Adrenal gland	Leukemia K-562	Fetal kidney	*E. coli* rRNA
D Parietal lobe	Amygdala	Pituitary gland	Atrium, right	Jejunum		Thymus	Uterus	Thyroid gland	Leukemia MOLT-4	Fetal liver	*E. coli* DNA
E Occipital lobe	Caudate nucleus	Spinal cord	Ventricle, left	Ileum		Peripheral blood leukocyte	Prostate	Salivary gland	Burkitt's lymphoma Daudi	Fetal spleen	Poly r(A)
F Temporal lobe	Hippocampus		Ventricle, right	Ilocecum		Lymph node	Testis	Mammary gland	Colorectal adenocarcinoma SW480	Fetal thymus	Human col-1 DNA
G Paracentral gyrus of cerebellum cortex	Medulla oblongata		Interventricular septum	Appendix		Bone marrow	Ovary		Lung carcinoma A549	Fetal lung	Human DNA 100 ng
H Pons	Putamen		Apex of the heart	Colon, ascending		Trachea					Human DNA 500 ng

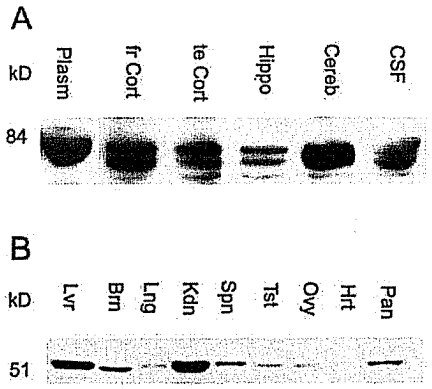

FIG. 4. **Tissue distribution of human CN1 and CN2.** *A*, 20 μg of extracts from human frontal cortex, temporal cortex, hippocampus, and cerebellum (Clontech), as well as 20 μg of CSF and plasma proteins, were immunoblotted on polyvinylidene difluoride membrane and probed with Y18K antibody. *B*, a commercial multiple tissue distribution blot (Oncogene) was used to analyze CN2 expression in human tissues. Blots were probed with the immunopurified anti-CN2 antibody (S17E). *Cereb*, cerebrellum; *Hippo*, hippocampus; *Te Cort*, temporal cortex; *Fr. Cort*, frontal cortex; *plasm*, plasma; *Lvr*, liver; *Brn*, brain; *Lng*, lung; *Kdn*, kidney; *Spn*, spleen; *Tst*, testis; *Ovy*, ovary; *Hrt*, heart; *Pan*, pancreas.

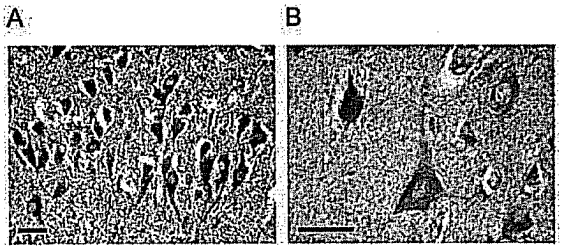

FIG. 5. **Immunohistochemical detection of CN1 protein in human brain.** Experiments were done as described under "Experimental Procedures" using antibody C17E. Signal is observed in hippocampal pyramidal neurons (*A*) and in the gray matter of the temporal cortex (*B*). Similar results were obtained with antibody K18K. *Scale bar*, 40 μm. Omission of the primary antibodies to CN1 protein or pre-incubation of antibody anti-C17E with peptide C17E resulted in no staining of the tissue sections (not shown).

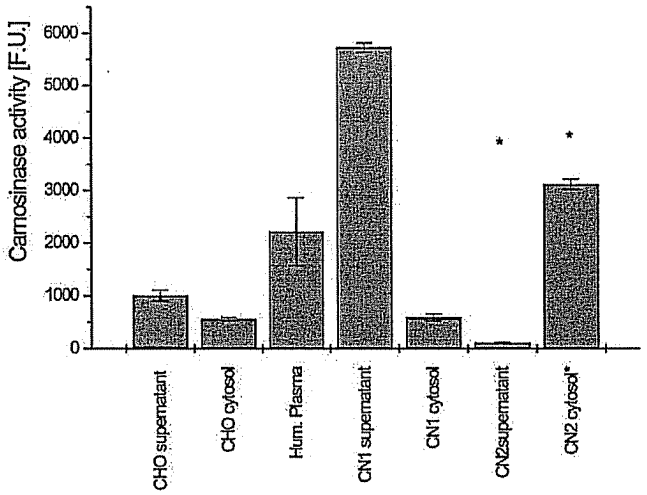

FIG. 6. **Detection of carnosinase activity in cells expressing CN1 or CN2.** Soluble cell extracts (0.5 μg) or cell culture supernatants (5 μl) from wild-type or cells transfected with CN1 or CN2 were assayed for carnosinase activity using conditions as described under "Experimental Procedures." *, samples were incubated in 50 mM Tris-sarcosine, pH 9.5, 0.1 mM $MnCl_2$.

transfected CHO cells (Fig. 7, *D* and *E*). Purified recombinant CN2 protein migrated at 54 kDa in SDS-PAGE, and the molecular mass determined by MALDI-TOF mass spectrometry was 52.8 kDa, which is in very good agreement to its predicted value deduced from the amino acid sequence (52.7 kDa). As deduced from size exclusion chromatography, the CN2 protein apparently forms a homodimer of 90 kDa.

Enzymatic Characterization and Substrate Specificity of CN1 and CN2—Taken together the experimental biophysical data for CN1 are in good agreement with literature data for human carnosinase (EC 3.4.13.20) (24). As for CN2, the cytoplasmic localization, the biophysical data (molecular mass, homodimer), and the capability of carnosine degradation at basic pH values are similar to those described for human tissue carnosinase (EC 3.4.13.18, later renamed cytosolic nonspecific dipeptidase) (39). To discriminate between both enzymes and further support this hypothesis both purified recombinant enzymes were characterized for their pH optimum of carnosine hydrolysis, substrate specificity, inhibition by a selection of protease inhibitors, and the effects of metal ions on enzyme activity. The pH activity curve of CN1 protein shows a rather broad maximum between pH 7.5 and 8.5 essentially as described for human carnosinase (24). For CN2 protein a narrow pH optimum for carnosine degradation with a maximum around pH 9.5 was found. The carnosine degrading activity of CN2 was affected by the presence of manganese. At pH 9.5 the addition of 0.1 mM Mn^{2+} leads to a 5-fold increase of enzyme activity, which is in agreement with human cytosolic nonspecific dipeptidase (19).

The substrate specificity of CN1 was determined with 17 Xaa-His dipeptides, three non Xaa-His dipeptides, and two His-containing tripeptides at pH 7.5 (Fig. 8*A*). Highest enzyme activity was found with carnosine (βAla-His), and only four of the other Xaa-His dipeptides served as substrate for this enzyme, including *N*-methylcarnosine, Ala-His, Gly-His, and GABA-His (homocarnosine). Contrary to the literature for human carnosinase, CN1 did not degrade anserine under the conditions used. The non Xaa-His dipeptides βAla-Ala, Ala-Ala, or Ala-Pro, as well as tripeptides containing histidine in central or C-terminal position (Gly-His-Gly or Gly-Gly-His), were not degraded, indicating that CN1 is a true Xaa-His dipeptidase. The catalytic efficiencies (k_{cat}/K_m) of CN1 for carnosine and homocarnosine were 8.6 and 1.1 $mM^{-1} \times s^{-1}$, respectively (Table II).

The substrate specificity of CN2 protein differs significantly. At pH 7.5 and in presence of 0.1 mM Mn^{2+}, the enzyme did not hydrolyze carnosine but degraded Pro-Ala and Ser-Gln with good activity. Whereas the Ser-Gln hydrolysis was optimal at pH 7.5, the velocity of Pro-Ala degradation was higher at pH 9.5. Only under these non-physiological conditions (pH 9.5) carnosine served as substrate for CN2, but as depicted in Fig. 8*B*, the relative dipeptidase activity with Xaa-His substrates like Leu-His, Ser-His, or Tyr-His was superior over carnosine degradation. Homocarnosine was not hydrolyzed at all, confirming previous data that human cytosolic nonspecific dipeptidase does not degrade this substrate (39). At pH 7.5, the k_{cat}/K_m of CN2 for Ser-Gln was at 21.5 $mM^{-1} \times s^{-1}$ with a K_m of 1.04 mM. For carnosine, the Michaelis constant of CN2 at pH 9.5 was 15 mM, which is in good agreement with literature data for human cytosolic nonspecific dipeptidase. Considering both the unphysiological pH optimum and the high K_m for carnosine, it appears that βAla-His is not a physiological substrate for this enzyme.

A set of commercially available protease inhibitors was selected to study inhibition of CN1 or CN2 activity. At 1 mM final concentration, inhibitors of serine, aspartic, and cysteine proteases (phosphoramidone, pepstatin, AEBSF, benzamide, E64,

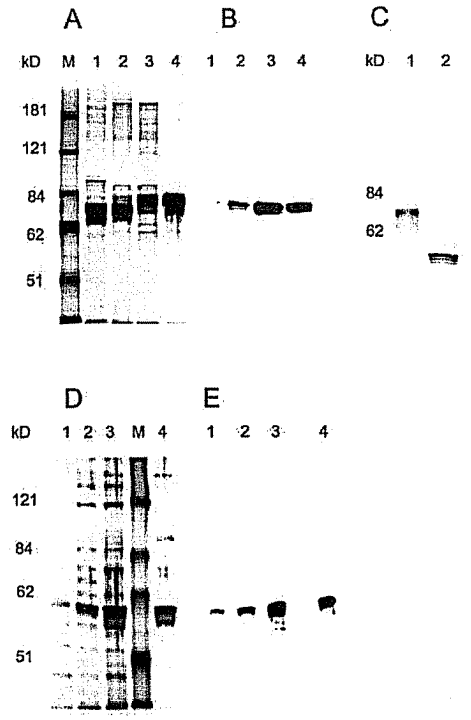

FIG. 7. **Purification of recombinant human CN1 and CN2.** *A*, analysis of purification steps of recombinant CN1 by SDS-PAGE. Recombinant CN1 was produced using a stable CHO cell line expressing the CN1 gene (*lane 1*) and purified using DEAE-Sephacel (*lane 2*) size exclusion chromatography (*lane 3*) and MonoQ anion exchange chromatography (*lane 4*). *Lane M*, marker proteins. 20 μg of proteins were loaded, separated on 8% SDS-PAGE, and silver-stained using a commercial kit (Invitrogen). *B*, detection of CN1 protein by Western blot. Protein samples were loaded as in *A* and transferred to a nitrocellulose membrane. Blots were analyzed using the immunopurified polyclonal C17E antibody and enhanced chemiluminescence detection. *C*, deglycosylation of purified recombinant CN1 by PNGaseF treatment. Purified carnosinase before PNGaseF treatment (*lane 1*) and deglycosylated carnosinase after PNGaseF treatment (*lane 2*) are shown. 100 μg of carnosinase were incubated for 18 h at 30 °C with 10 μg of PNGaseF before separation on SDS-PAGE and silver staining of the gels. *D*, analysis of purification steps of recombinant CN2 by SDS-PAGE. Recombinant CN2 was produced by transient transfection of CHO K1 cells with the cloned CN2 gene and purified from the soluble cell extract after 20000 × *g* centrifugation (*lane 1*) followed by DEAE-Sephacel (*lane 2*), size exclusion chromatography (*lane 3*), and MonoQ anion exchange chromatography (*lane 4*). *Lane M*, marker proteins. 20 μg of proteins were separated by 8% SDS-PAGE and silver-stained using a commercial kit (Invitrogen). *E*, detection of CN2 by Western blot. Protein samples from different purification steps (see *A*) were loaded on a 8% SDS gel and transferred to a nitrocellulose membrane. Blots were analyzed using the immunopurified polyclonal S16E antibody and enhanced chemiluminescence detection.

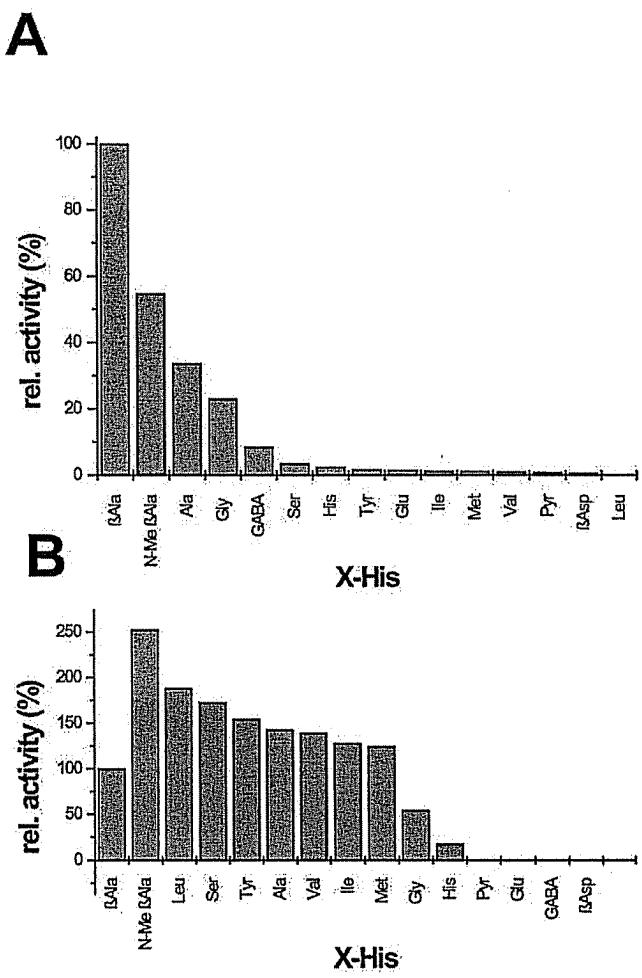

FIG. 8. **Substrate specificity of human CN1 and CN2 proteins for Xaa-His dipeptides.** Purified recombinant CN1 (*A*) or CN2 (*B*) proteins were incubated for 1 h at 30 °C with 17 Xaa-His dipeptides at 1 mM final concentration, and histidine release was detected by the fluorescent microplate assay (see "Experimental Procedures"). CN1 enzyme activity was measured after 1 h of incubation at 30 °C in 50 mM Tris-HCl, pH 7.5, using 10 ng of enzyme. CN2 activity was measured as for CN1 but using 50 mM Tris-sarcosine, pH 9.5, 0.1 mM $MnCl_2$. Values are expressed as relative activity setting the degradation of carnosine to 100%.

leupeptin) were active neither on CN1 nor on CN2 under standard assay conditions (see "Experimental Procedures"). The metal chelator 1,10-*o*-phenantrolin was inactive on CN2 but inhibited CN1 with an IC_{50} value of 5 μM. The cystein proteinase inhibitor *p*-hydroxymercurybenzoate inhibited CN2 with an IC_{50} of 13 μM whereas it was inactive on CN1.

The inhibition of CN1 and CN2 by bestatin, a compound known to specifically inhibit various amino- and dipeptidases (40), was analyzed in more detail. Recombinant CN2 was inhibited by bestatin with an IC_{50} of 7 nM, which is in good agreement with data on purified human cytosolic nonspecific dipeptidase. It should be noted that the assay was performed under non-physiological conditions (pH 9.5). Bestatin did not inhibit CN1 under our standard assay conditions (no additional metal ions) but showed an IC_{50} value of 12 μM in the presence of 0.2 mM Cd^{2+} as reported for human carnosinase (39).

TABLE II
Kinetic constants of human CN1 for carnosine or homocarnosine degradation and dependence of additional Cd^{2+} ions

Carnosine or homocarnosine breakdown was measured after 1 h of incubation in 50 mM Tris-HCl, pH 7.5, at 30 °C using 10 ng of purified recombinant CN1 protein preincubated for 15 min with or without 200 μM Cd^{2+} ions. The HPLC-based method was used to quantify release histidine (for details see "Experimental Procedures").

	Carnosine		Homocarnosine	
	$-Cd^{2+}$	$+Cd^{2+}$	$-Cd^{2+}$	$+Cd^{2+}$
K_m (mM)	1.27	11.0	0.2	1.9
k_{cat} (s^{-1})	10.6	138	0.2	18.04
k_{cat}/K_m ($mM^{-1} \cdot s^{-1}$)	8.6	12.2	1.1	9.5

The influence of additional Cd^{2+} on the activity of CN1 was studied in more detail by testing both the activity of CN1 and the inhibition by bestatin at different cadmium concentrations. Whereas at higher concentrations (>100 μM) cadmium activated CN1 protein, at low concentrations (0.1–3 μM) the metal ion strongly inhibited the enzyme (Fig. 9*B*). The saturation with Cd^{2+} ions had a strong effect on both the affinity and turnover of carnosine and homocarnosine degradation, because

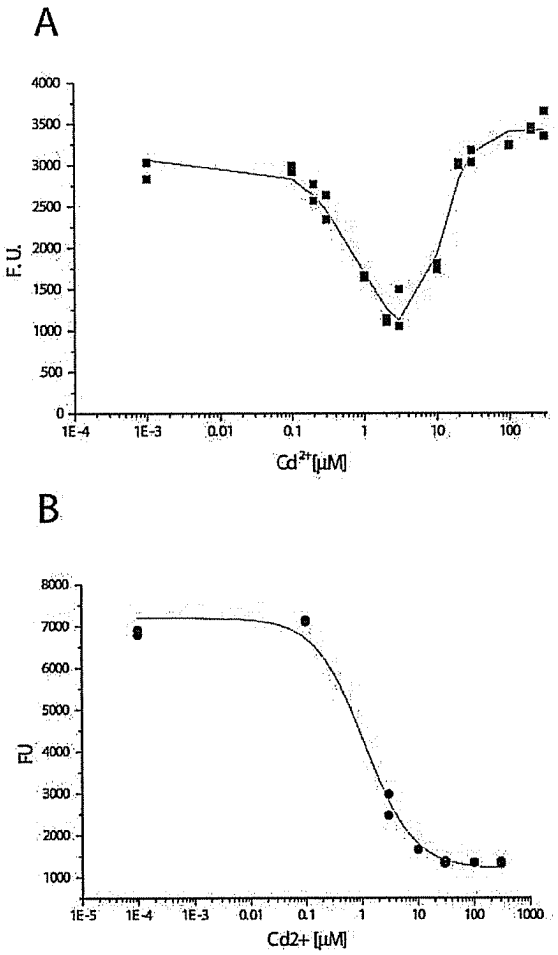

FIG. 9. **Influence of cadmium on human CN1 protein activity and inhibition by bestatin.** *A*, purified recombinant CN1 protein (10 ng) was incubated with 1 mM carnosine in the presence of different concentrations of cadmium for 1 h at 30 °C at pH 7.5. Enzyme activity was detected using the fluorescent microplate assay (see "Experimental Procedures"). The *curve* was calculated from a fit of the data using the equation, $v = (V_1/1 + [Cd]/K_i) + (V_2[Cd]^2/[Cd]^2 + K_a)$, where K_i and K_a are inhibition and activation constants, V_1 is the maximal velocity in the absence of Cd^{2+}, and V_2 is the velocity in the presence of saturating Cd^{2+}. When deriving the equation it was assumed that enzyme complexes with 1 Cd^{2+} are inactive. The calculated parameters are K_i = 1.2 ± 0.14 µM and K_a = 10.8 ± 0.7 µM. *B*, purified recombinant CN1 protein (10 ng) was incubated with 1 mM carnosine in the presence of different concentrations of cadmium in the presence of 50 µM bestatin (final concentration) for 1 h at 30 °C. Enzyme activity was detected using the fluorescent microplate assay (see "Experimental Procedures").

in the presence of 0.2 mM Cd^{2+}, K_m and k_{cat} are 10-fold higher for both substrates as compared with conditions where cadmium was omitted (Table II). As shown in Fig. 9*B* the inhibitory potency of the bestatin correlated with increasing concentrations of Cd^{2+}. Other metal ions (Fe^{2+}, Al^{3+}, Co^{2+}, Ni^{2+}) with the exception of Cu^{2+}, which strongly inhibited CN1 activity, had no effect on carnosine breakdown.

Functional Characterization of CN1 in the Human Neuroblastoma Cell Model—To assess the potential function of carnosinase in a physiological environment, CN1 was transiently overexpressed in human neuroblastoma cell line SH-SY5Y. Previous data showed that carnosine is capable of reversing the cytotoxic effect of MDA toward cultured rat brain endothelial cells (29). Our preliminary experiments showed that in the presence of 3 mM MDA, the viability of SH-SY5Y cells was decreased by 40%. Both carnosine (Fig. 10) and homocarnosine

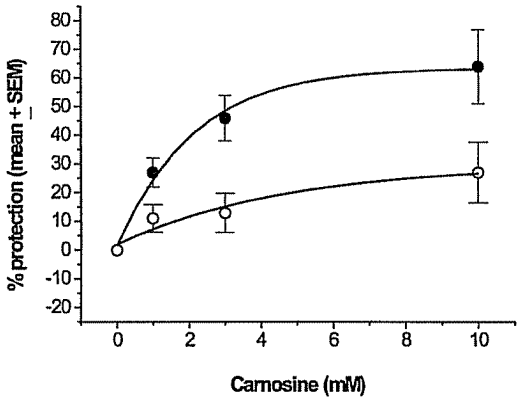

FIG. 10. **CN1 forms an active enzyme in human SH-SY5Y neuroblastoma cells.** Carnosine at 3 mM protects SHSY5Y neuroblastoma cells from MDA-induced toxicity (3 mM). Transient expression of human CN1 in SH-SY5Y cells (*open circles*) reduces capacity of carnosine to rescue the cells in comparison to MOCK transfected cells (*closed circles*). Results shown are the mean of five experiments.

(not shown) are capable of protecting neuroblastoma cells from MDA toxicity. At a concentration of 3 mM carnosine reverses the cytotoxic effect of MDA and protects 50% of control-transfected cells. This effect is reduced to 13% protection in cells transiently transfected with CN1 gene (Fig. 10). Transient expression of CN1 *per se* produced no toxic or proliferative effects on the cells. Additionally, the concentration of exogenous carnosine added to the medium was found to be reduced by 90%, most probably because of carnosinase activity (not shown). Collectively, these results clearly demonstrate that CN1 protein in SH-SY5Y cells forms a functional enzyme able to degrade carnosine and homocarnosine in a physiological environment.

DISCUSSION

In this study the identification, expression, and characterization of two human metal ion-dependent dipeptidases, CN1 and CN2, are described. Attracted by the brain-specific expression pattern of the CN1 transcript, the potential function of the CN1 protein and its widely expressed homologue CN2 was assessed by bioinformatic analysis. Using HMM and threading analysis the CN1 and CN2 polypeptides were found to be structurally related to proteins of the MEROPS M20 metalloprotease family. Both proteins expressed in CHO cells were purified to homogeneity and intensively characterized. Based on the distinct enzymatic properties, such as substrate specificity and protease inhibitor profile, biophysical data, and tissue distribution of both enzymes, we provide evidence that CN1 encodes the human carnosinase (EC 3.4.13.20), whereas CN2 is the cytosolic nonspecific dipeptidase (EC 3.4.13.18).

Our data show that both proteins are metal ion-dependent dipeptidases. Four point mutations introduced in CN1 completely abolished carnosinase activity, and metal ion chelating by 1,10-*o*-phenantrolin inhibited the enzyme in the low micromolar range. More precisely the results indicate that carnosinase and very likely cytosolic nonspecific dipeptidase use two metal ions in a co-catalytic mechanism. This hypothesis is supported by (i) bioinformatic analysis of the CN1 and CN2 sequences, which revealed a structural homology of both enzymes to the two-zinc enzyme carboxypeptidase G and (ii) kinetic data obtained for the inhibition and activation of carnosinase by cadmium. The selective activation of carnosinase by cadmium was described in the literature as a criterion to distinguish between the different carnosine-splitting enzymes and can be used as evidence for a two-metal ion co-catalytic mechanism. Although the intrinsically bound metal ions in

carnosinase are unknown, based on the homology to other enzymes of the M20 family of metalloproteases it can be assumed that carnosinase is associated with Mn^{2+} and/or Zn^{2+}. This would explain why the activity of carnosinase varies in the presence of cadmium. At low cadmium concentrations Mn^{2+}/Zn^{2+} in a binding site M1 could be replaced with this ion, and catalysis of carnosine degradation is inhibited, whereas enzyme activity is reconstituted if both metal binding sites (M1 and M2) are occupied by cadmium ions increasing K_m and k_{cat} for both carnosine and homocarnosine. Similar data were obtained for other co-catalytic metalloenzymes such as inositol monophosphatase (41) or fructose 1,6-bisphosphatase (42). Cytosolic nonspecific dipeptidase can also be activated by Mn^{2+}, Co^{2+}, and Cd^{2+}, which may explain why in previous studies (24, 44) human carnosinase (secreted form) could not be distinguished from the cytosolic nonspecific dipeptidase.

Although structurally related, CN1 and CN2 have quite different properties. Carnosinase is rather selective for Xaa-His dipeptides including homocarnosine, a brain-specific dipeptide (45), and carnosine. The cytosolic nonspecific dipeptidase degrades a large number of dipeptides. Our observation that βAla-His degradation by cytosolic nonspecific dipeptidase is observed only under non-physiological conditions supports the IUBMB recommendation to rename tissue carnosinase as nonspecific cytosolic dipeptidase (EC 3.8.13.18).

The biological function of both enzymes, although not yet well understood, may be quite different, as well. Whereas human CN1 RNA and protein are specifically expressed in adult central nervous system, CN2 RNA and protein expression are widely distributed in central and peripheral human tissues of adults. Interestingly, fetal brain does not contain CN1 mRNA suggesting that CN1 gene expression is induced with age. This result explains why carnosinase activity was virtually undetectable in newborn humans and increased with age (27). The origin of carnosinase in human plasma, however, still remains unknown. Rat and mouse orthologues of human carnosinase are found in the kidney but are not expressed in the CNS, supporting previous results describing a homocarnosine-splitting enzyme activity in the kidney of these animals (16). Given its broad tissue distribution, cellular localization, and broad substrate specificity the cytosolic nonspecific dipeptidase may function as a housekeeping enzyme in the catabolism of dipeptidic substrates. In contrast, it appears that carnosinase plays a special role in the homoeostasis of homocarnosine and carnosine in the human brain. Carnosine is known to quench destructive protein oxidizing agents deleterious to the cells. An even more important property of carnosine is its anti-advanced glycation end-products (AGE) activity. AGEs accumulate during aging, and the resulting cross-linking of protein deposits has been shown to occur in both plaques and tangles in Alzheimer's disease patients. Formation of AGEs is accelerated by transition metals, such as copper and iron. Carnosine is known to buffer free copper, to block formation of AGEs, and to inhibit protein cross-linking. Therefore, it may be important for cell viability to maintain the concentration of carnosine and probably of homocarnosine at high concentration. Literature data show that compared with younger individuals, older individuals exhibit significantly lower levels of free homocarnosine in CSF (43, 46).

In complement to the *in vitro* biochemical studies of the purified recombinant enzyme we showed that CN1 is functional in a physiological environment, particularly in the neuroblastoma cells SH-SY5Y. In this cell culture system carnosinase is able to metabolize its substrates added to the medium. This conclusion is based on the observation that SH-SY5Y cells transfected with CN1 were no longer protected against MDA-induced toxicity by carnosine and homocarnosine. In agreement with the literature, these data support the hypothesis that carnosine and homocarnosine are directly implicated in the protection of cells from toxicity of the products of lipid peroxidation and that carnosinase activity may be crucial for maintenance of protective activity of carnosine and homocarnosine in human brain. In conclusion, the identification of the long sought for carnosinase gene and the detailed characterization of the purified recombinant protein provide useful tools to study the biological role of this enzyme in aging and neurodegenerative or psychiatric diseases.

Acknowledgment—We thank F. Roux for technical assistance.

REFERENCES

1. Gulewitsch, W., and Amiradgibi, S. (1900) *Ber. Dtsch. Chem. Ges.* **33,** 1902–1903
2. Quinn, P., Boldyrev, A., and Formazuyk, V. (1992) *Mol. Aspects Med.* **13,** 379–444
3. Mannion, A. F., Jakeman, P. M., Dunnett, M., Harris, R. C., and Willan, P. L. (1992) *Eur. J. Appl. Physiol.* **64,** 47–50
4. Bale-Smith, E. C. (1938) *J. Phys.* **92,** 336–343
5. Abe, H. (2000) *Biochemistry (Mosc.)* **65,** 757–765
6. Tabakman, R., Lazarovici, P., and Kohen, R. (2002) *J. Neurosci. Res.* **68,** 463–469
7. Baran, E. J. (2000) *Biochemistry (Mosc.)* **65,** 789–797
8. Tromblay, P. Q., Horning, M. S., and Blakemore, L. J. (2000) *Biochemistry (Mosc.)* **65,** 807–816
9. McFarland, G., and Holliday, R. (1994) *Exp. Cell Res.* **167,** 175–184
10. Hipkiss, A. R. (1998) *Int. J. Biochem. Cell Biol.* **30,** 863–868
11. Petroff, O. A. C., Hyder, F., Rothman, D. L., and Mattson, R. H. (2001) *Neurology* **56,** 709–715
12. Petroff, O. A. C., Hyder, F., Rothman, D. L., and Mattson, R. H. (1999) *Neurology* **52,** 473–478
13. Horinishi, H., Grillo, M., and Margolis, F. L. (1978) *J. Neurochem.* **31,** 909–919
14. Hanson, H. T., and Smith, E. L. (1949) *J. Biol. Chem.* **179,** 789–801
15. Wood, T. (1957) *Nature* **180,** 39–40
16. Margolis, F. L., Grillo, M., Grannot-Reisfeld, N., and Farbman, A. I. (1983) *Biochim. Biophys. Acta* **744,** 237–248
17. Wolos, A., Piekarska, K., Głogowski, J., and Koieczka, I. (1978) *Int. J. Biochem.* **9,** 57–62
18. Kunze, N., Kleinkauf, H., and Bauer, K. (1986) *Eur. J. Biochem.* **160,** 605–613
19. Lenney, J. F., Peppers, S. C., Kucero-Ozallo, C. M., and George, R. P. (1985) *Biochem. J.* **228,** 653–660
20. Lenney, J. F. (1990) *Biol. Chem. Hoppe-Seyler* **371,** 167–171
21. Perry, T. L., Hansen, S., and Love, D. L. (1968) *Lancet* **8,** 1229–1230
22. Zoch, E., and Muller, H. (1971) *Enzymologica* **40,** 199–208
23. Lenney, J. F., George R. P., Weiss, A. M., Kucera, C. M., Chan P. W. H., and Rinzler, G. S. (1982) *Clin. Chim. Acta* **123,** 221–231
24. Jackson, M. C., Kucera, C. M., and Lenney, J. F. (1991) *Clin. Chim. Acta* **196,** 193–206
25. Clarke, R. (1994) in *Basic Cell Culture, A Practical Approach* (Davis, J. M., ed) pp. 223–242, IRL Press, Cambridge, MA
26. Gorman, C. M., Howard, B. H., and Reeves, R. (1983) *Nucleic Acids Res.* **11,** 7631–7648
27. Bando, K., Shimotsuji, T., Toyoshima, H., Hayashi, C., and Miyai, K. (1984) *Ann. Clin. Biochem.* **21,** 510–514
28. Lindroth, P., and Mopper, L. (1979) *Anal. Chem.* **51,** 1667–1674
29. Hipkiss, A. R., Preston, J. E., Himswoth, D. T. M., Worthington, V. C., and Abbot, N. J. (1997) *Neurosci. Lett.* **238,** 135–138
30. Kozak, M. (1989) *J. Cell Biol.* **108,** 229–241
31. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) *J. Mol. Biol.* **215,** 403–410
32. Eddy, S. R. (1998) *Bioinformatics* **14,** 755–763
33. Rawlings, N. D., and Barrett, A. J. (1995) *Methods Enzymol.* **248,** 183–228
34. Neuwald, A. F., Liu, J. S., Lipman, D. J., and Lawrence, C. E. (1997) *Nucleic Acids Res.* **25,** 1665–1677
35. Koppensteiner, W. A., Lackner, P., Wiederstein, M., and Sippl, M. (2000) *J. Mol. Biol.* **296,** 1139–1152
36. Rowsell, S., Pauptit, R. A., Tucker, A. D., Melton, R. G., Blow, D. M., and Brick, P. (1997) *Structure* **5,** 337–347
37. Nakai, K., and Horton, P. (1999) *Trends Biochem. Sci.* **24,** 34–35
38. Von Heijne, G. (1997) *Protein Eng.* **10,** 1–6
39. Peppers, S. C., and Lenney, J. (1988) *Biol. Chem. Hoppe-Seyler* **369,** 1281–1286
40. Suda, H., Aoyagi, H., Takeuchi, T., and Umezawa, H. (1976) *Arch. Biochem. Biophys.* **177,** 196–200
41. Pollack, S., Atack, J., Knowles, M., McAllister, G., Ragan, I., Baker, R., Fletcher, S., Iversen, L., and Broughton, H. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91,** 5766–5770
42. Zhang, Y., Liang, J. Y., Hunng, S., Ke, I. I., and Lipscomb, W. N. (1993) *Biochemistry* **32,** 1844–1857
43. Ferraro, T. N., and Hare, T. A. (1985) *Brain Res.* **338,** 53–60
44. Murphey, W. H., Patchen, L., and Lindmark, D. G. (1972) *Clin. Chim. Acta* **42,** 309–314
45. Kish, S. J., Perry, T. L., and Hansen, S. (1979) *J. Neurochem.* **32,** 1629–1636
46. Perry, T. L., Hansen, S., Stedma, D., and Love, D. (1968) *J. Neurochem.* **15,** 1203–1206